NPS55RH72111A

NAVAL POSTGRADUATE SCHOOL
Monterey, California

VALIDATION TESTS FOR

SHIPS SUPPLY SUPPORT STUDY

by

F. Russell Richards

November 1972

ved for public release; distribution unlimited

TABLE OF CONTENTS

NAVAL POSTGRADUATE SCHOOL
Monterey, California

Rear Admiral M. B. Freeman, USN
Superintendent

M. U. Clauser
Provost

ABSTRACT:

A study is devoted to the development of statistical procedures to be used to test the validity of the Ships Supply Support Study simulator. Some theory is presented for each test procedure, but special emphasis is paid to describing each test in detail so that the tests can be implemented by the project study group. Examples which illustrate the required numerical work are given with each test procedure, and the advantages and disadvantages of each are pointed out. A computer program which will perform the calculations necessary for the more complicated test is included.

This test was supported by the Research and Development Division, Naval Supply Systems Command, under NAVSUP RDT&E No. TF 38.531.001.

Prepared by:

1. INTRODUCTION

A problem associated with any simulation study is to test the results obtained from the simulator to determine how well the system being studied is duplicated. It is only after the simulation results have been validated that any reliance can be placed on the output of that simulator. In the Ships Supply Support Study (S^4) a simulator is used to obtain estimates of the mean supply response time (MSRT) for each of 13 material cognizance classes (cogs). In addition, random samples of actual response times observed by ships in the fleet for each of the cogs have been collected. In order to validate the S^4 simulation results the simulator MSRT's should be compared to the available data. Since cog response times are random variables with perhaps different distributions for each cog, statistical methods are desired to test whether or not the simulator MSRT's are indeed the same as the means of the actual cognizance class response times, the populations in this study.

There are several well known statistical techniques which can be used to perform the above tests. This report describes briefly some of those tests and discusses some of the advantages and disadvantages of each. Detailed instructions for implementing the tests are presented, and examples are given to illustrate the numerical work which must be done.

Although the statistical tests are developed and described with reference to S^4, they should also be useful for validating future simulations.

2. NOTATION

The following standard notation will be used in describing the statistical tests in this report. Let n_i be the number of actual ship observations of supply response time for cog i. Then, let X_{ij} be the j^{th} observation of supply response time for cog i; $i = 1, 2, \ldots, 13$ and $j = 1, 2, \ldots, n_i$. Let

$$\bar{X}_i = \sum_{j=1}^{n_i} X_{ij} / n_i$$

and

$$S_i^2 = \sum_{j=1}^{n_i} (X_{ij}-\bar{X}_i)^2 / (n_i-1),$$

the sample means and sample variances, respectively, for cog i. The value μ_i will denote the actual (unknown) population mean supply response time for cog i, and λ_i will represent the value obtained from the S^4 simulator for the MSRT of cog i.

3. A SINGLE TEST FOR THE MULTIPLE HYPOTHESIS TESTING PROBLEM

3.1 BACKGROUND

It is desirable to validate the S^4 simulator by testing whether or not each simulator MSRT approximates the population MSRT. In this section, a single test is derived which will validate or invalidate the entire set of simulator MSRT's. Thus, it offers a computational advantage of requiring few calculations.

Consider the hypothesis testing problem:

$$H_o: \mu_i = \lambda_i \qquad i = 1, 2, \ldots, 13$$

vs

$$H_1: \mu_i \neq \lambda_i \qquad \text{for some } i.$$

H_o is the null hypothesis, and it simply states that each of the simulator MSRT's is the actual mean of the population of the respective cog response time. The alternative hypothesis, H_1 states that some simulator MSRT is not the population MSRT. If each of the sample sizes $n_1, n_2, \ldots, n_{13}$ is sufficiently large (50 or more), this hypothesis testing problem does not require any assumptions about the underlying populations of the supply response times other than requiring that the populations have finite first and second moments. This is important because little work has been done in determining the properties of the response time distributions.

3.2 DERIVATION OF THE TEST STATISTIC

With fairly large sample sizes the central limit theorem indicates that the distribution of each $\overline{X}_i$ is approximately normal with mean μ_i and variance σ_i^2 / n_i, regardless of the population distribution. Thus,

$$Z_i^2 = \left(\frac{(\overline{X}_i - \mu_i)}{\sigma_i / \sqrt{n_i}} \right)^2$$

is approximately chi-square distributed with one degree of freedom, denoted $\chi^2(1)$. Furthermore, the $\overline{X}_i$'s are independent so that

$$Y = \sum_{i=1}^{13} Z_i^2$$

is chi-square distributed with 13 degrees of freedom, $\chi^2(13)$. For the statistical problem at hand the population variance σ_i^2, like the population mean, is unknown. To circumvent this problem

the sample variance S_i^2 is substituted for the population variance σ_i^2. This substitution should not be serious if the sample sizes are large.

When the null hypothesis is true; that is, $\mu_i = \lambda_i$ for all i, the distribution of

$$D = \sum_{i=1}^{13} \frac{(\bar{X}_i - \lambda_i)^2 n_i}{S_i^2}$$

will be approximately chi-square with 13 degrees of freedom. On calculating a value of D based on the observations in a random sample, the value should not be large if the null hypothesis is true. To be more precise, let $\chi_\alpha^2(13)$ represent that unique number such that the probability that an observation of a random variable having a distribution which is chi-square with 13 degrees of freedom exceeds $\chi_\alpha^2(13)$ is α. Then, if the null hypothesis is true, an observed value of D would exceed the value $\chi_\alpha^2(13)$ with a probability of α. If α is small, it is an unlikely event that a value of D exceeds $\chi_\alpha^2(13)$. The statistician is led to believe that H_o is false whenever such an event occurs.

3.3 DESCRIPTION OF THE TEST PROCEDURE

The test procedure reduces to an especially simple comparison of the calculated value of D with the number $\chi_\alpha^2(13)$.

(1) Determine the maximum probability that one is willing to reject H_o when H_o is in fact true. Denote this probability, the level of significance of the test, by α.

(2) Determine the number $\chi^2_\alpha(13)$ from standard chi-square tables.

(3) Calculate a value of D based on the observed random samples.

(4) If the value of D exceeds the number $\chi^2_\alpha(13)$ reject H_o.

3.4 DETERMINING THE POWER OF THE STATISTICAL TEST

When the level of significance, α, is chosen the probability of rejecting the null hypothesis when it is true is determined. It is also useful to determine the probability of rejecting the null hypothesis when it is false, the power of the test. Suppose $\mu_i = \theta_i$ and for at least one value of i, $\theta_i \neq \lambda_i$. The random variable D is now distributed as a non-central chi-square random variable with 13 d.f. and non-centrality parameter

$$\beta = \frac{1}{2} \sum_{i=1}^{13} (\theta_i - \lambda_i)^2 .$$

Tables for the non-central chi-square distribution can be used to determine the probability of rejecting H_o when the alternative $\mu_i = \theta_i$; for $i = 1, 2, \ldots, 13$ is true. The test procedure is known to have the desirable property that the probability of rejecting H_o when it is false is greater than α; that is, H_o is more likely to be rejected when it is false than when it is true.

3.5 EXAMPLE

For brevity, suppose that the number of cogs is three and let the following table give the values determined from the random

samples and the simulator.

i	$\bar{X}_i$	S_i^2	n_i	λ_i
1	20	16	48	21
2	32	4	48	32
3	24	16	48	26

TABLE 1

The hypothesis being tested is:

$$H_o: \ \mu_1 = \lambda_1, \ \mu_2 = \lambda_2, \ \text{and} \ \mu_3 = \lambda_3$$

vs

$$H_1: \ \mu_1 \neq \lambda_1, \ \text{or} \ \mu_2 \neq \lambda_2 \ \text{or} \ \mu_3 \neq \lambda_3$$

Let the level of significance be .05. The number $\chi^2_{.05}(3)$ is found from chi-square tables to be 7.815. The value of the test statistic D calculated from the given values is:

$$D = \frac{48(-1)^2}{16} + \frac{48(0)}{4} + \frac{48(-2)^2}{16} = 15$$

Since the value of the test statistic D exceeds the value $\chi^2_{.05}(3)$ the null hypothesis is rejected at a .05 level of significance.

3.6 DISADVANTAGES OF TEST 1

The example points out two disadvantages of this test. When the test statistic rejects the null hypothesis, nothing is revealed about the validity of the individual simulator values. It is probably true that some of the simulator values are sufficiently accurate even when the null hypothesis is rejected. In fact, the data in the above example shows that $\bar{X}_2 = \lambda_2$ and S_2^2 is small, so that the simulator value for cog 2 is obviously good. It seems desirable

that the statistical procedure not only test the hypothesis, but also locate those cogs on which further work need be done.

Another disadvantage revealed by the example is that each of the simulator values is within ten percent of the sample means, but the null hypothesis is rejected. Since a ten percent error is probably a tolerable amount as far as validating the simulator, perhaps the hypothesis should be stated as:

$$H_o: \quad (1-p)\, \lambda_i \le \mu_i \le (1+p)\, \lambda_i \qquad i = 1, 2, \ldots, 13$$

vs

$$H_1: \quad \mu_i < (1-p)\, \lambda_i \quad \text{or} \quad \mu_i > (1+p)\, \lambda_i \qquad \text{some } i$$

where $0 \le p \le 1$ is a permissible fraction of error. The second statistical test considers such a null hypothesis.

4. INDIVIDUAL STATISTICAL TESTS FOR COMPOSITE TWO-SIDED HYPOTHESES.

4.1 BACKGROUND

Two disadvantages of the first test procedure were pointed out. The test developed in this section eliminates both of those disadvantages. Consider testing each cog, individually. A test statistic will be derived for each cog, and individual values of cog MSRT are either validated or invalidated. Let k be any one of the 13 cogs. The statistical hypothesis tested in this section is:

$$H_o: \qquad \theta_1 \le \mu_k \le \theta_2$$

vs

$$H_1: \qquad \mu_k < \theta_1 \quad \text{or} \quad \mu_k > \theta_2.$$

For S^4 applications θ_1 and θ_2 would probably be taken to be $(1-p)\,\lambda_k$ and $(1+p)\,\lambda_k$, respectively. This tests the hypothesis that the true mean differs from the simulated mean by no more than a fraction p.

Although attention is restricted to a single cog at a time, the fact that the statistical hypotheses are two-sided and composite complicates the problem. Nevertheless, a test can be derived which has "optimal" properties. Lehmann [2] has shown that the test which will be developed is the uniformly most powerful unbiased test of size α. It is by this criterion that the test is considered best.

4.2 DESCRIPTION OF THE TEST

Since the distributions of the populations are unknown, it is necessary to assume that the sample sizes are sufficiently large so that the sample variances can be substituted for the unknown population variances in the test procedure. The large sample sizes are also needed to justify appealing to the central limit theorem to take the distribution of $\overline{X}_k$ to be normal with mean μ_k and variance σ_k^2 / n_k.

Let R_k be that subset of the sample space containing all of those points $(X_1, X_2, \ldots, X_n)$ such that their sample mean $\overline{X}$ is either greater than a constant L_k or less than a constant U_k; i.e.,

$$R_k = \{(X_1, \ldots, X_n) \mid \overline{X} < L_k \text{ or } \overline{X} > U_k\}.$$

The uniformly most powerful unbiased test of size α has a particularly simple form with the set R_k as the rejection region. Thus, if $(X_{k1}, X_{k2}, \ldots, X_{kn_k})$, the vector of sample response times for cog k,

is in the set R_k reject H_o. Otherwise, accept H_o. Hence, all that remains to specify completely the best test is to find the values of the constants L_k and U_k. These are determined from the conditions which state that the probabilities of rejecting H_o when $\mu_k = (1-p)\lambda_k$ and $\mu_k = (1+p)\lambda_k$ must both be α in order that the test be unbiased of level α. That is,

$$P[\bar{X}_k < L_k \text{ or } \bar{X}_k > U_k \mid \mu_k = (1-p)\lambda_k] = \alpha$$

and

4.2.1 $$P[\bar{X}_k < L_k \text{ or } \bar{X}_k > U_k \mid \mu_k = (1+p)\lambda_k] = \alpha$$

Since the distribution of $\bar{X}_k$ is symmetric and μ_k is a location parameter for that distribution, it is easily shown that $L_k = \lambda_k - C_k$ and $U_k = \lambda_k + C_k$. Because the problem reduces to finding the single constant C_k only one of the probability statements is necessary. Thus, for a given level α equation 4.2.1 can be used to uniquely determine C_k.

By standardizing the random variable $\bar{X}_k$ (using $S_k/\sqrt{n_k}$ in place of $\sigma_k/\sqrt{n_k}$) equation 4.2.1 can be replaced by the equivalent statement:

4.2.2 $$P\left[\frac{-C_k - p\lambda_k}{S_k/\sqrt{n_k}} \le Z \le \frac{C_k - p\lambda_k}{S_k/\sqrt{n_k}}\right] = 1 - \alpha,$$

where Z is standard normal random variable, $N(0,1)$. Let $F(y) = P[Z \le y]$, $a_k = \sqrt{n_k}\, p\lambda_k / S_k$ and $b_k = \sqrt{n_k}\, C_k / S_k$. To determine the constant C_k or equivalently the constant b_k, the zero of the equation:

4.2.3 $$F(b_k - a_k) - F(-b_k - a_k) = 1 - \alpha$$

must be determined. Some one dimensional search procedure is needed to find that value since an explicit solution is unknown. Fortunately, in many cases the search is unnecessary for $F(-b_k-a_k)$ is approximately zero. When this is the case, standard normal distribution tables can be used to look up the value of b_k such that

4.2.4 $$F(b_k-a_k) = 1 - \alpha.$$

4.3 OUTLINE OF PROCEDURE

1. Choose a level α and a fraction p and calculate $\bar{X}_k$, S_k, and a_k for each cog k.
2. Using the standard normal distribution tables find the value of b_k such that $F(b_k-a_k) = 1 - \alpha$.
3. Calculate $F(-b_k-a_k)$. If this value is sufficiently small (.0005 should be small enough), take b_k to be the value determined in step 2 and proceed to step 6.
4. If $F(-b_k-a_k) > 0.0005$, choose a new value $b_k' > b_k$.
5. If the value of $F(b_k'-a_k) - F(-b_k-a_k) > 1 - \alpha$ decrease b_k' and try again. If the value is less than $1 - \alpha$ increase b_k' and repeat this step.
6. When the value of $F(b_k'-a_k) - F(-b_k'-a_k) = 1 - \alpha$, C_k is then $b_k' S_k / \sqrt{n_k}$, $L_k = \lambda_k - C_k$ and $U_k = \lambda_k + C_k$.
7. If $\bar{X}_k > U_k$ or $\bar{X}_k < L_k$ reject H_o. Otherwise accept H_o.

4.4 EXAMPLES

A. 1. Test the hypothesis that the mean supply response time for cog 1 is within 10% of the simulator MSRT.

$$H_o: \quad 0.90\,\lambda_1 \le \mu_1 \le 1.10\,\lambda_1$$

vs

$$H_1: \quad \mu_1 < 0.90\,\lambda_1 \quad \text{or} \quad \mu_1 > 1.1\,\lambda_1$$

Let the level of significance be $\alpha = .05$. The sample mean and the sample standard deviation based on a random sample of size 100 are found to be $\overline{X}_1 = 38$ and $S_1 = 5$. The simulator MSRT is $\lambda_1 = 35$ and a_1 is calculated to be 7.

2. From the standard normal tables find $b_1 - 7 = 1.645$ and thus $b_1 = 8.645$.

3. Since $b_1 + a_1 = 15.645$ no iterative search need be made; $C_1 = 4.3225$, $L_1 = 30.6775$ and $U_1 = 39.3225$.

4. Since $\overline{X}_1$ is in the interval [30.6775, 39.3225] the null hypothesis is not rejected.

B. 1. Test the same hypothesis as above, but suppose the mean and standard deviation calculated from a random sample of size 36 are $\overline{X}_1 = 38$ and $S_1 = 35$. Let the simulator MSRT still be $\lambda_1 = 35$ and calculate a_1 to be 0.60.

2. From standard normal tables find $b_1 - 0.60 = 1.645$, and thus $b_1 = 2.245$.

3. Since $F(-b_1-a_1) = .00225$, choose a new value $b_1' > b_1$. Suppose the new value is $b_1' = 2.280$.
4. The value of $F(b_1'-a_1) - F(-b_1'-a_1) = F(1.680) - F(-2.880) = 0.9535 - 0.0020 = 0.9515$ which is now larger than $1 - \alpha$. Choose a new value $b_1' < 2.280$, say $b_1' = 2.265$.
5. The value of $F(b_1'-a_1) - F(-b_1'-a_1)$ is now $0.9520 - 0.0021 = .9499$, or within .0001 of the value of $1 - \alpha$, and the iterative search is terminated.
6. Calculate $C_1 = 13.21$, $L_1 = 21.79$ and $U_1 = 48.21$.
7. Since $\bar{X}_1$ is in the interval [21.79, 48.21] the null hypothesis is not rejected.

Examples A and B illustrate the simple procedures which must be followed in order to test the hypothesis that the true population mean is within a given percentage of the simulator MSRT's. For the data which is to be tested in the Ships Supply Support Study it would probably rarely be necessary to perform an iterative search for the constants C_k. In example B, the sample standard deviation was forced to be unrealistically large in order to present a case where a search was necessary. Even if a search were necessary, a trial and error procedure by hand quickly zeros in on the solution.

4.5 AN EXTENSION OF TEST 2

Whenever the null hypothesis stating that the population mean supply response time is within a given percentage of the simulator MSRT is rejected, it seems desirable to know how poor is the simulator value. The test considered in this section can be used to provide additional information

when the hypothesis is rejected. Instead of stating only that the hypothesis is rejected at a given level for a given value of p, we can also determine the smallest value of p for which the hypothesis would be accepted at the given level. For example, it may be the case that the hypothesis stating that the population mean supply response time is within 10% of the simulator MSRT at the .05 level is rejected, but the hypothesis would be accepted at the same level if p were 0.15. The magnitude of the minimum value of p needed to accept the null hypothesis should provide some information as to the amount of additional work needed, if any, to be satisfied that the simulator is generating reasonable numbers.

4.6 DETERMINING THE MINIMUM VALUE OF p

We have seen that the null hypothesis is accepted at a given level α provided that $\lambda_k - C_k \leq \overline{X}_k \leq \lambda_k + C_k$, where C_k depends on the value of p. We use this to determine the minimum value of p needed in order that the acceptance interval contain the number $\overline{X}_k$.

1. Take C_k to be the absolute value of the difference between λ_k and $\overline{X}_k$; that is,
$$C_k = | \lambda_k - \overline{X}_k |.$$
2. Determine $b_k = C_k \sqrt{n_k}/S_k$.
3. Since p is unknown, a_k must be determined iteratively. As an initial trial take $a_k = b_k - Z_\alpha$.
4. Calculate $F(-b_k-a_k)$. If this value is, say, less than 0.0005, the minimum value of p is
$$p_{min} = \frac{a_k S_k}{\sqrt{n_k}\, \lambda_k} .$$

If the value of $F(-b_k-a_k)$ is too large to be neglected, the value of a_k must be determined by using an iterative search just as the value of b_k was determined previously.

5. If $F(-b_k-a_k) > 0.0005$, choose a new value $a_k' < a_k$.
6. If $F(b_k-a_k') - F(-b_k-a_k') < 1 - \alpha$ decrease a_k' and try again. If the value is greater than $1 - \alpha$ increase a_k' and repeat this step.
7. When a_k' is such that $F(b_k-a_k') - F(-b_k-a_k') = 1 - \alpha$,

$$p_{min} = \frac{a_k' \, S_k}{\lambda_k \sqrt{n_k}} .$$

The following example illustrates the above calculations. Let us find the minimum value of p which will cause the null hypothesis

$$H_o: \lambda_1 - p\lambda_1 \leq \mu_1 \leq \lambda_1 + p\lambda_1$$

to be accepted at a .05 level of significance where $\lambda_1 = 35$. The sample mean and sample standard deviation based on a sample of size 100 are $\overline{X}_1 = 41$ and $S_1 = 5$. We find that $C_1 = | 41 - 35 | = 6$ and $b_1 = C_1 \sqrt{n_1} / S_1 = 12$. For the initial trial we take $a_1 = b_1 - Z_\alpha = 12.0 - 1.645 = 10.355$. Since $-b_1-a_1 = -22.355$, $F(-b_1-a_1)$ is negligible, and we take a_1 to be 10.355. The minimum value of p for which the null hypothesis is accepted is therefore

$$p_{min} = \frac{(10.355)(5)}{(10)(35)} = 0.1479.$$

As with the search for b_k in the hypothesis testing problem with the given value of p an iterative search will rarely be necessary in the above procedure. In almost all problems of practical interest a_k will be approximately $b_k - Z_\alpha$ since $F(-a_k-b_k)$ will be negligible.

4.7 THE COMPUTER PROGRAM

Although the calculations needed to carry out the tests in this section are quite simple and a trial and error search usually zeros in very rapidly on the values of b_k or a_k, it may not be feasible to perform a large number of tests manually. For that reason a computer program utilizing a first-order Newton search, whenever a search is necessary, is included as Appendix A. The program only requires the user to input the number of tests, a value for p, the significance level, and a tolerable error which is used to terminate the search. (An error of 0.0005 was used in the examples.) In addition, the user must input the sample size, the sample mean, the sample standard deviation, and the simulator MSRT for each cog. The comment cards in the program should make the program self explanatory.

For each cog the program determines the lower and upper limits of the acceptance region for the given value of p, and it tells if the hypothesis is accepted or rejected. If an hypothesis is rejected the program gives the minimum value of p necessary to accept the hypothesis. Included with the program in Appendix A is a set of sample input and sample output.

5. SUMMARY AND CONCLUSIONS

Two tests have been investigated for use in testing the validity of the S^4 simulator. The single assumption necessary for the procedures to be valid seems easily justified based on the sample data which was available. Both of the tests presented are quite simple and could easily be performed by the group at FMSO or some other member of the S^4 team in only a few hours.

Because of the disadvantages of Test 1 which were pointed out in Section 3.5, Test 2 is recommended. Theoretically, it is the best test possible for testing the given hypothesis, and the included computer program will perform a large number of tests in a few seconds. A caution should be pointed out concerning the choice of p for Test 2. When choosing a value for p, one should be aware that he is essentially declaring a tolerable error for the simulation values. It is necessary that the effects of such errors on any conclusions arrived at by S^4 be examined. If the effects are severe p must be small.

It is the author's intent that this report be useful as a working paper for performing the needed tests for S^4 and perhaps also for future studies.

APPENDIX A

```
C**************************************************************C
C                                                              C
C    VALIDATION TESTS FOR SSSS SIMULATOR                       C
C    FIRST-ORDER NEWTON SEARCH IS USED WHEN NEEDED             C
C                                                              C
C**************************************************************C
C                                                              C
C  DEFINITION OF VARIABLES                                     C
C                                                              C
C**************************************************************C
C                                                              C
C    NTEST       NUMBER OF TESTS TO BE MADE                    C
C    ALPHA       SIGNIFICANCE LEVEL OF THE TESTS               C
C    P           FRACTION OF ERROR FOR THE TESTS               C
C    ZALPHA      VALUE FROM NORMAL TABLES SUCH THAT            C
C                P( Z > ZALPHA ) = ALPHA                       C
C    ERROR       ERROR TOLERANCE.  THE ITERATIVE SEARCH        C
C                WILL TERMINATE ONLY IF                        C
C                ABS(F(B-A) - F(-B-A) - (1-ALPHA)) < ERROR     C
C    AN(I)       SIZE OF SAMPLE I                              C
C    XBAR(I)     MEAN OF SAMPLE I                              C
C    S(I)        STANDARD DEVIATION OF SAMPLE I                C
C    SMEAN(I)    SIMULATOR MEAN FOR COG I                      C
C    PMIN        MINIMUM VALUE OF P FOR WHICH HYPOTHESIS       C
C                IS ACCEPTED                                   C
C    L           LOWER LIMIT OF THE ACCEPTANCE REGION          C
C    U           UPPER LIMIT OF THE ACCEPTANCE REGION          C
C                                                              C
C**************************************************************C
      DOUBLE PRECISION IDEC1,IDEC2
      DATA IDEC1/' ACCEPT '/,IDEC2/' REJECT '/
      DIMENSION AN(20), S(20), XBAR(20), SMEAN(20)
C**************************************************************C
C                                                              C
C    READ INPUT                                                C
C  READ NO. OF TESTS, ALPHA, P, ZALPHA, AND ERROR              C
C  READ SAMPLE SIZE,XBAR,STD. DEV., AND SIMULATOR MEAN FOR     C
C  EACH TEST.                                                  C
C                                                              C
C**************************************************************C
      READ(5,1) NTEST,ALPHA,P,ZALPHA,ERROR
    1 FORMAT(I5,F5.3,3F10.5)
      READ(5,3) (AN(I),XBAR(I),S(I),SMEAN(I), I=1,NTEST)
    3 FORMAT(8F10.0)
C**************************************************************C
C                                                              C
C    SET UP OUTPUT TABLE                                       C
C                                                              C
C**************************************************************C
      WRITE(6,55)  ALPHA, P
   55 FORMAT('0',6X,'***** V A L I D A T I O N   T E S T S *'
     1,'****',///,'     HYPOTHESIS TESTED: (1-P)LAMBDA < MU'
     2,' < (1+P)LAMBDA',//,5X,'LEVEL OF SIGNIFICANCE= ',
     3F5.3,5X,'P=',F6.3,///)
      WRITE(6,50)
   50 FORMAT(3X,58(1H-))
      WRITE(6,51)
   51 FORMAT(3X,'|  N | DECISION |    L     |     U     |',
     1'   XBAR    |  PMIN |',/,3X,58(1H-))
C**************************************************************C
C                                                              C
C    BEGIN CALCULATIONS FOR TESTS 1 THROUGH NTEST              C
C                                                              C
C**************************************************************C
      DO 40 K=1,NTEST
      ROOTN = SQRT(AN(K))
      A = P * SMEAN(K) * ROOTN / S(K)
      B = ZALPHA + A
```

```
      IFLAG = 0
    2 X1 = B - A
      X2 = -B - A
      CALL NDTR(X1,CDF1,PDF1)
      IF(X2.GE. -4.0) GO TO 23
      CDF2 = 0.0
      PDF2 = 0.0
      GO TO 33
   23 CALL NDTR(X2,CDF2,PDF2)
   33 FX = CDF1 - CDF2 - (1.0 - ALPHA)
      IF(IFLAG.EQ.0) GO TO 34
C***********************************************************************C
C                                                                       C
C     CAN WE TERMINATE THE SEARCH ON A?                                 C
C                                                                       C
C***********************************************************************C
      IF(ABS(FX).LT.ERROR) GO TO 17
C***********************************************************************C
C                                                                       C
C     CONTINUE ITERATIVE SEARCH                                         C
C                                                                       C
C***********************************************************************C
      FPRIME = PDF2 - PDF1
      A = A - FX/FPRIME
      GO TO 2
C***********************************************************************C
C                                                                       C
C     CAN WE TERMINATE THE SEARCH ON B?                                 C
C                                                                       C
C***********************************************************************C
   34 IF(ABS(FX).LT.ERROR) GO TO 7
C***********************************************************************C
C                                                                       C
C     CONTINUE ITERATIVE SEARCH                                         C
C                                                                       C
C***********************************************************************C
      FPRIME = PDF1 + PDF2
      B = B - FX/FPRIME
      GO TO 2
C***********************************************************************C
C                                                                       C
C     DETERMINE ACCEPTANCE INTERVAL                                     C
C                                                                       C
C***********************************************************************C
    7 C = B * S(K) / ROOTN
      XMAX = SMEAN(K) + C
      XMIN = SMEAN(K) - C
C***********************************************************************C
C                                                                       C
C     IS XBAR INSIDE THE ACCEPTANCE INTERVAL?                           C
C                                                                       C
C***********************************************************************C
      IF(XBAR(K).GT.XMAX.OR.XBAR(K).LT.XMIN) GO TO 9
C***********************************************************************C
C                                                                       C
C     YES, ACCEPT HYPOTHESIS AND WRITE ANSWERS                          C
C                                                                       C
C***********************************************************************C
      WRITE(6,6)  K,IDEC1, XMIN, XMAX, XBAR(K)
    6 FORMAT(3X,'|',I3,' | ',A8,' |',  F9.3,' |',F9.3,' |',
     1F9.3,' |          |')
      GO TO 40
C***********************************************************************C
C                                                                       C
C     NO, REJECT HYPOTHESIS AND DETERMINE PMIN                          C
C                                                                       C
C***********************************************************************C
    9 C = ABS(SMEAN(K) - XBAR(K))
      B = C * ROOTN / S(K)
      A = B - ZALPHA
      IFLAG = 1
      GO TO 2
```

```
C***********************************************************C
C                                                           C
C    CALCULATE PMIN AND WRITE ANSWERS                       C
C                                                           C
C***********************************************************C
   17 PMIN    = A * S(K) / (ROOTN * SMEAN(K))
      WRITE(6,18)   K, IDEC2, XMIN, XMAX, XBAR(K), PMIN
   18 FORMAT(3X,'|',I3,' | ',A8,' |',  F9.3,' |',F9.3,' |',
     1F9.3,' |',F6.3,' |')
   40 CONTINUE
      WRITE(6,50)
      STOP
      END

C***********************************************************C
C                                                           C
C    SUBROUTINE TO CALCULATE CDF AND PDF FROM NORMAL DIST.  C
C                                                           C
C***********************************************************C

      SUBROUTINE NDTR(X,P,D)
      AX=ABS(X)
      T=1.0/(1.0+.2316419*AX)
      D=0.3989423*EXP(-X*X/2.0)
      P= 1.0 - D*T*((((1.330274*T - 1.821256)*T + 1.781478)
     1 *T - 0.3565638)*T + 0.3193815)
      IF(X) 1,2,2
    1 P=1.0-P
    2 RETURN
      END
C***********************************************************C
C                                                           C
C    DATA                                                   C
C                                                           C
C***********************************************************C
# DATA
    6 0.05  0.1      1.645       0.0001
100.0      48.55      12.32       36.00       75.0        63.21
 12.0      35.6       64.33       11.33       49.0        36.0
 147.      42.0       9.3         57.0        196.        48.
```

```
C   VALIDATION TESTS FOR SSSS SIMULATOR.
C
C   USES FIRST ORDER NEWTON SEARCH WHEN SEARCH IS NEEDED
C
      DOUBLE PRECISION IDEC1,IDEC2
      DATA IDEC1/' ACCEPT '/,IDEC2/' REJECT '/
      DIMENSION AN(20), S(20), XBAR(20), SMEAN(20)
C
C   READ NO. OF TESTS, ALPHA, P, ZALPHA, AND ERROR
C
      READ(5,1) NTEST,ALPHA,P,ZALPHA,ERROR
    1 FORMAT(I5,F5.3,3F10.5)
C
C   READ SAMPLE SIZE,XBAR,STD. DEV., AND SIMULATOR MEAN FOR
C   EACH TEST.
      READ(5,3) (AN(I),XBAR(I),S(I),SMEAN(I), I=1,NTEST)
    3 FORMAT(8F10.0)
      WRITE(6,55)  ALPHA, P
   55 FORMAT('0',6X,'***** V A L I D A T I O N  T E S T S *'
     1,'****',///,'      HYPOTHESIS TESTED: (1-P)LAMBDA < MU'
     2,' < (1+P)LAMBDA',//,5X,'LEVEL OF SIGNIFICANCE= ',
     3F5.3,5X,'P=',F6.3,///)
      WRITE(6,50)
   50 FORMAT(3X,58(1H-))
      WRITE(6,51)
   51 FORMAT(3X,'|  N | DECISION |     L      |     U     |     X
     1'  PMIN |',/,3X,58(1H-))
      DO 40 K=1,NTEST
      ROOTN = SQRT(AN(K))
      A = P * SMEAN(K) * ROOTN / S(K)
      B = ZALPHA + A
      IFLAG = 0
    2 X1 = B - A
      X2 = -B - A
      CALL NDTR(X1,CDF1,PDF1)
      IF(X2.GE. -4.0) GO TO 23
      CDF2 = 0.0
      PDF2 = 0.0
      GO TO 33
   23 CALL NDTR(X2,CDF2,PDF2)
   33 FX = CDF1 - CDF2 - (1.0 - ALPHA)
      IF(IFLAG.EQ.0) GO TO 34
      IF(ABS(FX).LT.ERROR) GO TO 17
      FPRIME = PDF2 - PDF1
      A = A - FX/FPRIME
      GO TO 2
   34 IF(ABS(FX).LT.ERROR) GO TO 7
      FPRIME = PDF1 + PDF2
      B = B - FX/FPRIME
      GO TO 2
    7 C = B * S(K) / ROOTN
      XMAX = SMEAN(K) + C
      XMIN = SMEAN(K) - C
      IF(XBAR(K).GT.XMAX.OR.XBAR(K).LT.XMIN) GO TO 9
      WRITE(6,6)  K, IDEC1, XMIN, XMAX, XBAR(K)
    6 FORMAT(3X,'|',I3,' | ',A8,' |',  F9.3,' |',F9.3,' |',
     1F9.3,' |         |')
      GO TO 40
    9 C = ABS(SMEAN(K) - XBAR(K))
      B = C * ROOTN / S(K)
      A = B - ZALPHA
      IFLAG = 1
      GO TO 2
   17 PMIN    = A * S(K) / (ROOTN * SMEAN(K))
      WRITE(6,18)  K, IDEC2, XMIN, XMAX, XBAR(K), PMIN
   18 FORMAT(3X,'|',I3,' | ',A8,' |',  F9.3,' |',F9.3,' |',
     1F9.3,' |',F6.3,' |')
   40 CONTINUE
      WRITE(6,50)
      STOP
      END
```

```
      SUBROUTINE NDTR(X,P,D)
      AX=ABS(X)
      T=1.0/(1.0+.2316419*AX)
      D=0.3989423*EXP(-X*X/2.0)
      P= 1.0 - D*T*((((1.330274*T - 1.821256)*T + 1.781478)
     1 *T - 0.3565638)*T + 0.3193815)
      IF(X) 1,2,2
    1 P=1.0-P
    2 RETURN
      END
# DATA
    6 0.05   0.1       1.645      0.0001
100.0     48.55      12.32      36.00       75.0       63.21
 12.0     35.6       64.33      11.33       49.0       36.0
 147.     42.0       9.3        57.0        196.       48.
```

***** V A L I D A T I O N   T E S T S *****

HYPOTHESIS TESTED: (1-P)LAMBDA < MU < (1+P)LAMBDA
LEVEL OF SIGNIFICANCE = 0.050   P = 0.100

N	DECISION	L	U	XBAR	PMIN
1	REJECT	30.373	41.627	48.550	0.292
2	ACCEPT	54.222	70.438	63.210	
3	ACCEPT	-25.135	47.795	35.600	
4	REJECT	18.390	25.610	36.000	0.572
5	REJECT	50.038	63.962	42.000	0.241
6	ACCEPT	40.648	51.352	48.000	

BIBLIOGRAPHY

[1] Ferguson, Thomas S., Mathematical Statistics A Decision Theoretic Approach, Academic Press, 1967.

[2] Lehmann, E. L., Testing Statistical Hypotheses, John Wiley & Sons, Inc., 1959.

[3] Fix, E., "Tables of Non-central Chi-square," Univ. of Calif. Pub. in Statist., Vol. 1, 1949, pp. 15-19.

[4] Mood, Alexander M. and Graybill, F. A., Introduction to the Theory of Statistics, McGraw-Hill, 1963.

[5] Ships Supply Support Study Group, Second Interim Report, August, 1972.

INITIAL DISTRIBUTION LIST

	No. of Copies
Defense Documentation Center (DDC) Cameron Station Alexandria, Virginia 22314	12
Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
Library, Code 55 Naval Postgraduate School Monterey, California 93940	2
Dean of Research Administration (Code 023) Naval Postgraduate School Monterey, California 93940	1
Mr. Herbert J. Lieberman NSUP 01A1B Naval Supply Systems Command Washington, D. C. 20390	2
Mr. James W. Pritchard Code NSUP - 01A Naval Supply Systems Command Washington, D. C. 20390	1
LCDR James A. Biggins, SC, USN Fleet Material Support Office Mechanicsburg, Pennsylvania 17055	1
Professor F. Russell Richards (Code 55Rh) Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	10

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

REPORT TITLE

Validation Tests for the Ships Supply Support Study

. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report

. AUTHOR(S) (First name, middle initial, last name)

F. Russell Richards

. REPORT DATE November 1972	7a. TOTAL NO. OF PAGES 29	7b. NO. OF REFS 5

a. CONTRACT OR GRANT NO. NAVSUP RDT&E b. PROJECT NO. TF 38.531.001 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) NPS55RH72111A 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

0. DISTRIBUTION STATEMENT

Approved for public release, distribution unlimited

1. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Research and Development Division Naval Supply Systems Command

3. ABSTRACT

A study is devoted to the development of statistical procedures to be used to test the validity of the Ships Supply Support Study simulator. Some theory is presented for each test procedure, but special emphasis is paid to describing each test in detail so that the tests can be implemented by the project study group. Examples which illustrate the required numerical work are given with each test procedure, and the advantages and disadvantages of each are pointed out. A computer program which will perform the calculations necessary for the more complicated test is included.

This test was supported by the Research and Development Division, Naval Supply Systems Command, under NAVSUP RDT&E No. TF 38.531.001.

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
SHIPS SUPPLY SUPPORT STUDY						
SIMULATION						
VALIDATION TESTS						
STATISTICAL HYPOTHESIS TESTING						
MEAN SUPPLY RESPONSE TIME						